

January 16, 2025

VIA ECF

Hon. Ona T. Wang
United States District Court
500 Pearl Street
New York, NY 10007

Re: *The New York Times Company v. Microsoft Corporation, et al.*, 1:23-cv-11195; *Daily News, LP, et al. v. Microsoft Corporation, et al.*, 1:24-cv-3285

Dear Magistrate Judge Wang:

The News Plaintiffs' letter motion seeks an order requiring OpenAI to preserve all user conversation outputs, even if the conversation has no relationship to Plaintiffs' content, and even if a user has requested the conversation's deletion. This is an extraordinary request, especially given OpenAI's preservation of [REDACTED] ChatGPT user conversations. Declaration of Michael Trinh ("Trinh Decl.") ¶ 4. Plaintiffs' request for the preservation of irrelevant information is unduly burdensome and disproportionate to the needs of this case. Accordingly, this Court should deny Plaintiffs' letter motion.

OpenAI has more than 300 million weekly active users. Trinh Decl. ¶ 2. Consistent with OpenAI's Privacy Policy, OpenAI retains ChatGPT user conversation data by default. *Id.* ¶ 4. As described in the Privacy Policy, ChatGPT users may, subject to narrow exceptions, use in-product options to opt out of having their ChatGPT conversation data retained. *Id.* ¶ 5.

Between when Plaintiff The New York Times filed its original complaint on December 27, 2023, and now, more than [REDACTED] new ChatGPT user accounts have been created. *Id.* ¶ 7. Those accounts alone have engaged in more than [REDACTED] conversations over the last 13 months. *Id.* Across all users and all time, OpenAI currently has retained more than [REDACTED] ChatGPT conversations. *Id.* ¶ 4. For purposes of this litigation, OpenAI has further retained data associated with specific ChatGPT user accounts likely to have responsive information. Declaration of Joseph C. Gratz ("Gratz Decl."), Ex. A (Jan. 3, 2025 Letter from J. Gratz). In light of this extraordinary volume of data, counsel have for months engaged in productive and ongoing conversations regarding how to sample and use it.

Apparently not satisfied with sampling [REDACTED] preserved conversations for litigation analysis, Plaintiffs now request OpenAI "preserve [all of] its output log data" going forward, for all users and all conversations worldwide. *See* ECF No. 379 ("Mot.") at 1, 3. The Court should reject this request to preserve all user conversation outputs because (1) it targets irrelevant data, (2) it would be unduly burdensome, and (3) the request is untimely. Such a sweeping change to OpenAI's ordinary operations is disproportionate to the needs of the case.

Irrelevant. This motion is the first time OpenAI has understood Plaintiffs to be requesting the retention of *all* user conversation data, regardless of the relevance of each conversation to Plaintiffs or their content. In their December 10 correspondence, Plaintiffs made a more limited request, asking OpenAI to “(1) identify any output containing content from any of the News Plaintiffs that has been deleted; and (2) confirm in writing . . . that you will hereafter preserve ChatGPT user conversations containing any of the News Plaintiffs’ content.” ECF No. 379-15. In response, OpenAI has worked diligently over the holidays to evaluate the feasibility of Plaintiffs’ request. Trinh Decl. ¶ 8. OpenAI told Plaintiffs in a January 3, 2025 letter that it remained “willing to discuss preserving discrete subsets of” user conversation data. Gratz Decl., Ex. A (January 3, 2025 Letter from J. Gratz). During that ongoing dialogue, Plaintiffs focused on user conversations that implicate OpenAI’s retrieval-augmented generation or copyright guardrail systems, which Plaintiffs contended was likely to reference their content. ECF No. 379-15. The Court should reject Plaintiffs’ newfound, extraordinarily broader request for OpenAI to retain *all* user conversation output data.

Unduly Burdensome. Even as to conversation data subject to user-deletion requests which might arguably be relevant, Plaintiffs’ request is unduly burdensome. OpenAI’s Privacy Policy reflects the company’s commitment to protecting users’ privacy, including by reminding users that they may have the right to “delete your personal data from our records.” Trinh Decl. ¶ 3; *See Privacy Policy*, OpenAI (Updated Nov. 4, 2024) available at <https://openai.com/policies/row-privacy-policy/> (accessed Jan. 16, 2025). Because OpenAI’s default data retention policy respects user deletion requests, it would harm OpenAI’s reputation and customer relationships to override those practices across the board. Trinh Decl. ¶ 11. Such a burden is further unnecessary because of the cumulative nature of Plaintiffs’ request. Rather than require OpenAI to override its users’ privacy preferences, the better course is for counsel to continue to work together to develop a sampling methodology for the more than [REDACTED] user conversations already preserved.

Moreover, OpenAI does not have the preexisting technological functionality to fulfill Plaintiffs’ request to somehow preserve only user conversations that reference Plaintiffs’ content (to the extent all of Plaintiffs’ content could even be systematically identified. Trinh Decl. ¶ 9. Based on its diligent exploration of technical feasibility in response to Plaintiffs December 10 letter, *id.* ¶ 8, OpenAI currently anticipates that overriding user conversation deletion requests in order to search and then preserve only those conversations referencing Plaintiffs’ content, if even possible, would require a new custom-designed system. *Id.* ¶ 9.

Preliminary estimates from OpenAI’s privacy engineering team are that attempting to build such a system would require [REDACTED]

Id. ¶ 9.

[REDACTED] *Id.* Additional work would then be required to achieve a minimum viable concept, followed by committing a non-trivial ongoing amount of reliability maintenance engineering work to the system. *Id.* OpenAI is unable to reliably estimate this amount of maintenance work without a more detailed specification. *Id.*

Reallocating full-time engineering resources to attempt to build a custom solution to override user privacy preferences would be incredibly burdensome, and would impair OpenAI’s other ongoing initiatives and operations. *Id.* ¶ 10.

Untimely. Plaintiffs are incorrect that they only “became aware” of OpenAI’s user conversation data retention policy on November 15, 2024. Mot. at 2. Rather, OpenAI first told Plaintiffs about its policy nearly nine months earlier. On February 29, 2024, before OpenAI had even served its initial disclosures, OpenAI proactively wrote counsel for the New York Times and disclosed that it was “retaining input and output data from ChatGPT unless users have affirmatively turned off conversation history, deleted their past conversations, or deleted their account, in which case new conversations are retained for 30 days, as stated on OpenAI’s website.” ECF No. 379-3 at 1. More recently, OpenAI reminded Plaintiffs of those retention policies. ECF No. 379-13. Concurrently, counsel for OpenAI and Plaintiffs have engaged in ongoing discussions about a feasible means of producing relevant user conversation data. *See* Gratz Decl., Ex. A at 2 (“OpenAI is …collaborating with Plaintiffs to develop an approach to Plaintiffs’ requests regarding using that data in this litigation”). Because Plaintiffs have long been on notice of OpenAI’s user conversation data retention policies, their request is untimely and should be denied.

* * *

At the end of the day, this consolidated copyright litigation brought by the news-industry plaintiffs should not require OpenAI to undertake the drastic measure of disregarding user chat deletion requests in order to “preserve all ChatGPT user conversations.” Mot. at 3. Under the Federal Rules, “[l]itigants are not generally required to preserve all electronic records, and the principles of proportionality guide the scope of a party’s duty to preserve potentially relevant evidence.” *Doe LS 340 v. Uber Techs., Inc.*, 710 F. Supp. 3d 794, 804 (N.D. Cal. 2024) (denying motion to compel suspension of all document destruction policies at the company as “exceedingly broad”). Accordingly, OpenAI need not “preserve every shred of paper, every email or electronic document” since such a requirement could “cripple” its operations. *Zubulake v. UBS Warburg, LLC*, 220 F.R.D. 212, 217 (S.D.N.Y. 2003); *see also FTC v. DirecTV, Inc.*, No. 15-cv-01129-HSG (MEJ), 2016 WL 7386133, at *5 (N.D. Cal. Dec. 21, 2016) (a party need not preserve data beyond what is “needed . . . to prosecute th[e] case”). Considering the massive quantity of ChatGPT user conversation data already retained, the burden to the company of custom-engineering a solution (if even possible), and Plaintiffs’ delay in raising the issue, this Court should deny Plaintiffs’ motion.

January 16, 2025

Respectfully Submitted,

MORRISON & FOERSTER LLP
/s/ Joseph C. Gratz
 Joseph C. Gratz

KEKER, VAN NEST & PETERS LLP
/s/ Edward A. Bayley
 Edward A. Bayley

LATHAM & WATKINS LLP
/s/ Elana Nightingale Dawson
 Elana Nightingale Dawson